

# Fault tolerant Multi Cluster head Data Aggregation Protocol in WSN (FMCDCA)

Sushruta Mishra<sup>1</sup>, Lambodar Jena<sup>2</sup>, Alok Chakrabarty<sup>3</sup>, Jyotirmayee Choudhury<sup>4</sup>  
Department of Computer Science & Engineering<sup>1, 2, 3, 4</sup>,  
Gandhi Engineering College<sup>1, 2</sup>, NIT Meghalaya<sup>3</sup>, Essel Mining<sup>4</sup>  
Bhubaneswar<sup>1, 2</sup>, Meghalaya<sup>3</sup>, Barbil<sup>4</sup>, India

**Abstract**—WSN consisting of a large number of small sensors with low-power transceivers can be an effective tool for gathering data in a variety of environments. As sensor nodes are deployed in sensing field, they can help people to monitor and aggregate data. Researchers also try to find more efficient ways of utilizing limited energy of sensor node in order to give longer life time of WSNs. Network lifetime, scalability, and load balancing are important requirements for many data gathering sensor network applications. Therefore, many protocols are introduced for better performance. The efficient node-energy utilization is one of important performance factors in wireless sensor networks but besides this reliability is also an issue. In this paper, we proposed a cluster based routing algorithm to ensure high reliability such that the entire network becomes fault tolerant by the introduction of multi cluster heads.

**Keywords**- Wireless Sensor Networks; LEACH; Cluster-Based Routing; Ccluster Formation.

## I. INTRODUCTION

Wireless sensor network (WSN) consists of hundreds and even thousands of small tiny devices called sensor nodes distributed autonomously to monitor physical or environmental conditions, such as temperature, sound, vibration, pressure and motion at different locations. Energy plays an important role in wireless sensor networks because nodes are battery operated. Consequently many protocols have been proposed in order to minimize the energy consumption of these nodes. Each node in a sensor network is typically equipped with one or more sensors, a radio transceiver or other wireless communications device, a small microcontroller, and an energy source, since in most Wireless sensor network applications the energy source is a battery, energy plays an important role in wireless sensor network, and preserving the consumed energy of each node is an important goal that must be considered when developing a routing protocol for wireless sensor networks. A wireless sensor network is typically made of many sensor nodes used to detect accuracy and scalability of sensing areas. In such a large scale networking environment, one of the most important networking factors is the self-organizing capability for adaptation to dynamic situation changes and interoperating capability between sensor nodes [1]. Many studies have shown that there are a variety of sensors used for gathering sensing information and efficiently transferring the information to the sink nodes. Sensor networks need protocols, which are specific,

data centric, capable of aggregating data and optimizing energy consumption. An ideal sensor network should have the following additional features:

1. *Attribute based addressing* is typically employed in sensor networks. The attribute-based addresses are composed of a series of attribute-value pairs, which specify certain physical parameters to be sensed.
2. *Location awareness* is another important issue. Since most data collection is based on location, it is desirable that the nodes know their position whenever needed.

The major issues stemming from these studies are protocol design in regards to battery energy efficiency, localization scheme, synchronization, data aggregation and security technologies for wireless sensor networks. In particular, researchers have shown great interest in the routing protocols in the network layer, which considers self organization capabilities, limited battery power, and data aggregation schemes[2][3]. The routing protocol of sensor networks is typically partitioned into two sub routings: (1) flat routing protocol and (2) hierarchical routing protocol. The sensor node performs a data aggregation process to avoid duplicated data transfers. Such a sequence of processes favors the hierarchical routing protocol based upon clusters due to the fact that efficient selection of cluster heads can reduce the usage of consumption power and maximize the life time of the networks [4].

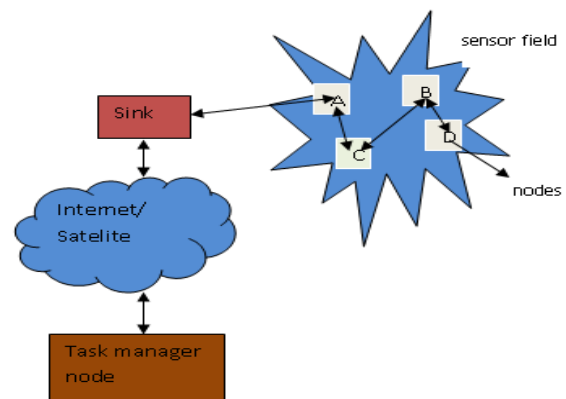


Figure 1. A simple wireless sensor network

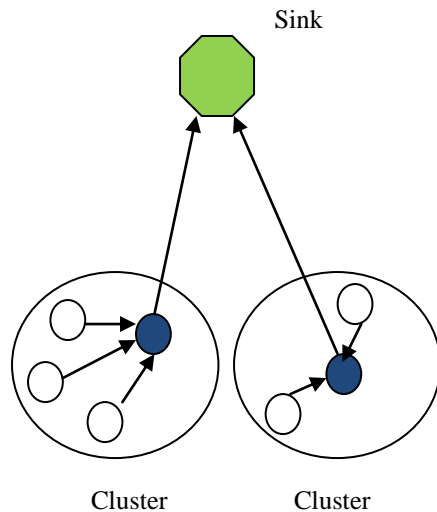


Figure 2. A cluster based model

#### A. Cluster based routing:

The basic objective on any routing protocol is to make the network useful and efficient. A cluster based routing protocol groups sensor nodes where each group of nodes has a CH or a gateway. Sensed data is sent to the CH rather than send it to the BS, CH performs some aggregation function on data it receives then send it to the BS where these data is needed. A number of routing protocols have been proposed for WSN. However, few of them are cluster based. Two of the most well known hierarchical protocols are LEACH, PAMAS and PEGASIS [6]. Both of these show significant reduction in the overall network energy over other non-clustering protocol. Hierarchical routing protocols designed to reduce energy consumption by localizing communication within the cluster and aggregate data to reduce transmissions to the BS.

#### B. Clustering objectives

There are different objectives of the clustering algorithms. The clustering objective is often established to fulfill applications requirements such as low data latency or data location awareness. Some of the popular objectives are discussed.

#### C. Load balancing

The distribution of sensors among clusters in an evenly manner is a common goal where cluster heads perform data processing or a significant amount of tasks. Load balancing is a more pressing issue in WSNs where cluster heads are chosen from available sensor nodes, since it becomes crucial to avoid the exhaustion of cluster heads prematurely.

#### D. Fault-tolerance

Tolerating the failure of cluster heads is usually necessary in applications where WSN are operating in harsh environments in order to avoid the loss of important Initialization algorithms for wireless ad-hoc networks sensor's data. Assigning backup cluster heads is the most notable scheme pursued in the literature for recovery from a cluster head failure. Rotating the role of cluster heads among nodes in the cluster can also be a means for fault-tolerance besides to their load balancing advantage.

## II. INCREASED CONNECTIVITY AND REDUCED DELAY

In WSNs, which the cluster heads are picked from sensor nodes, limiting the range of connectivity and enhancing inter-cluster heads connectivity may be more suitable than long-haul connections. On the other hand, when data latency is a concern, intracluster connectivity becomes a design objective or constraint.

#### A. Minimal cluster count

This is a common objective when cluster heads are specialized resource-rich nodes. In these cases their deployment is more difficult or tend to be more expensive and vulnerable than sensors.

#### B. Maximal network longevity

Since sensor nodes are constraint, the network's lifetime is a major concern especially for applications of WSNs placed in harsh environments. Adaptive clustering is a viable choice in order to achieve more network longevity.

## III. RELATED WORK

There have been many studies presented for configuration and operation algorithms of a cluster based network topology in ad hoc networks. Advantages to a clustering network are the reduction of an overhead of routing establishment, minimization of the size of the routing table, and stabilization of the network topology. The clustering network can make resource management and bandwidth allocation more efficient and make node positioning management and transmitting power management possible. In the clustering algorithm, all nodes in a sensor network can become a cluster head but must belong to only one cluster. The algorithm should minimize the overhead of clustering setup messages and establishing times. Additionally, the algorithm must maintain a stable network configuration, routing, network efficiency, with a minimization of energy consumption [4]. Many clustering algorithms have been proposed, most of which are based upon node identifier, node connectivity, and node weights. Some of better-known cluster based hierarchical routing protocol are LEACH, LEACH-Centralized, E-LEACH, TL-LEACH, M-LEACH and PEGASIS (Power Efficiency Gathering in Sensor Information Systems).

#### A. LEACH

LEACH [5] is a clustering routing protocol in which a cluster head collects data from sensor nodes belonging to the cluster and sends the data to the sink node after data aggregation process. To make all sensor nodes in this network consume their node energy equally and extend the life time of the network, this algorithm randomly changes the

cluster head, which in turn uses more energy than any other node belong to the cluster, every time period. To reduce overall communication costs, the cluster head performs data aggregation and then send the data to the sink node. The cluster head is determined by the following function (1):

$$T(n) = \begin{cases} \frac{p}{1-p * (r \bmod \frac{1}{p})} & \text{if } n \in G \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Where P is the desired percentage of cluster heads, r is the current round number, G is the set of nodes that have not been cluster-heads in the last 1/P rounds. A round consists of two phases; a set-up phase and a steady state phase. The former is a stage for configuration of a cluster head and a cluster, and the latter is a stage for data transfer by the TDMA schedule.

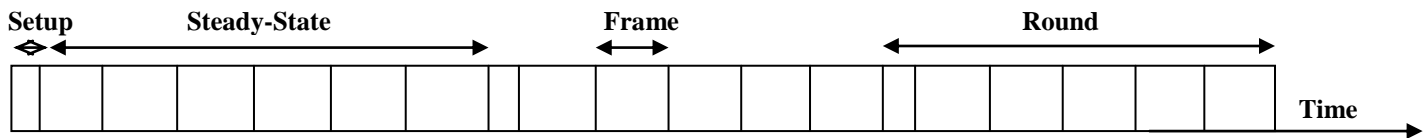


Figure 3. Time line of LEACH algorithm

When a new round starts, each sensor node generates a random number in the range of 0 and 1, computes a threshold value by using equation (1), and compares the two numbers. If the generated number is smaller than the threshold value, the node is nominated as a cluster head; otherwise it neglects the number and remains a plain node. The nominated cluster head broadcasts advertisement messages over neighbor nodes. The neighbor node that receives the advertisement messages selects one of broadcasting nodes that transmits the strongest broadcasting signal as its head cluster node, and sends a "Join-REQ" message to the head cluster. After receiving the "Join-REQ" message, the head cluster registers the node onto its own member node table. The cluster head makes a TDMA schedule for data transfer within the cluster network and broadcasts the schedule to its member nodes. It is at this point that the setup phase to select a cluster head has completed. In the next steady state phase, each node in a cluster network sends information data to its cluster head by the TDMA schedule. The cluster head send the aggregated data to the sink node, called its base station. To reduce the overhead of the cluster head selection once a cluster head has been selected, many rounds of data frame transfer are performed followed by a repeat of the cluster reconfiguration procedure. Since LEACH uses a probability in selection of cluster heads, its advantage is that all nodes have a chance to becoming a cluster head. But the stochastic process brings unbalance of node energy consumption which ultimately shortens the life time of the sensor network.

#### B. LEACH-C (LEACH – Centralized)

As previously mentioned, the disadvantage to LEACH is that the number of cluster head nodes is little ambiguous to count. LEACH-C [6] has been proposed to clarify this problem. LEACH-C provides an efficient clustering configuration algorithm, in which an optimum cluster head is selected with minimization of data transmission energy between a cluster head and other nodes in a cluster. In LEACH-C, the base station receives information about residual node energy and node positions at the set up phase of each round. The received data can compute an average residual energy for all nodes. The nodes with less than average energy are excluded in selection of cluster

heads. Among the nodes that have more than average energy, cluster heads are selected with use of the simulated annealing algorithm. The base station sends all nodes a message of the optimum cluster head IDs (Identifiers). The node, the ID of which is the same as the optimum cluster head ID, is nominated as a cluster head and prepares a TDMA schedule for data transfer. Other nodes wait for the TDMA schedule from their cluster heads.

#### C. E-LEACH Protocol

Energy-LEACH protocol improves the CH selection procedure. It makes residual energy of node as the main metric which decides whether the nodes turn into CH or not after the first round [10].

Same as LEACH protocol, E-LEACH is divided into rounds, in the first round, every node has the same probability to turn into CH, that mean nodes are randomly selected as CHs, in the next rounds, the residual energy of each node is different after one round communication and taken into account for the selection of the CHs. That mean nodes have more energy will become a CHs rather than nodes with less energy.

#### D. TL-LEACH

In LEACH protocol, the CH collects and aggregates data from sensors in its own cluster and passes the information to the BS directly. CH might be located far away from the BS, so it uses most of its energy for transmitting and because it is always on it will die faster than other nodes. A new version of LEACH called Two-level Leach was proposed. In this protocol; CH collects data from other cluster members as original LEACH, but rather than transfer data to the BS directly, it uses one of the CHs that lies between the CH and the BS as a relay station [8].

#### E. M-LEACH protocol

In LEACH, Each CH directly communicates with BS no matter the distance between CH and BS. It will consume lot of its energy if the distance is far. On the other hand, Multihop-LEACH protocol selects optimal path between the CH and the BS through other CHs and use these CHs as a relay station to transmit data over through them [9]. First, multi-hop communication is adopted among CHs. Then, according to the selected optimal path, these CHs transmit data to the corresponding CH which is nearest to BS. Finally, this CH sends data to BS. M-LEACH protocol is almost the same as LEACH protocol, only makes communication mode from single hop to multi-hop between CHs and BS.

#### F. PEGASIS

The chain formation in PEGASIS[6][7] leads to a high latency as all the data has to pass through the chain to reach the base station, if the farthest or the first node of the chain has important information which has to be passed immediately will have to travel through the entire chain. It can easily add the newly deployed nodes in the chain as it is not a fixed transmitting path. If it finds a new node which saves much more energy it adds it during the chain formation. It has high energy awareness due to formation of chain structure to reach the base station which is much more energy conserving that cluster formation in LEACH. A low overhead is seen on the network as there are no other nodes which transmit other than all nodes that form the chain and only one node is responsible which is near to the destination or the sink to transmit. The quality of service factor is low as there is a delay in the data transmission and no processing capabilities, all nodes fuse some data with the data packet while forwarding to other nodes in the chain. Network instability like a node failure or link failure or power failure can cause loss of data.

## IV. PROPOSED WORK

In the following, we will describe the deployment and method of the protocol. According to the above mentioned

routing protocols, the network assumptions can be initiated as follows [4, 5, 6].

1. Each node or sink has ability to transmit message to any other node and sink directly.
2. Each sensor node has radio power control node can tune the magnitude according to the transmission distance.
3. Each sensor node has the same initial power in WSNs.
4. Each sensor node has location information.
5. Every sensor nodes are fixed after they were deployed.
6. WSNs would not be maintained by humans.
7. Every sensor nodes have the same process and communication ability in WSNs, and they play the same role.
8. Wireless sensor nodes are deployed densely and randomly in sensor field.

We call our proposed model as FMCDA (Fault tolerant Multi Cluster head Data Aggregation) There are three major issues in our proposed algorithm and they are as follows :

1. The sink node is the main controller of the entire operation
2. The selection of cluster heads are based on the residual energies on the sensor nodes after every round
3. A multi cluster head scheme is implemented where two cluster heads are present in a cluster. (One PCH and one SCH)

The idea of multi cluster heads is based on the fact that if the main cluster head fails then the other one can form the backup up so that the cluster do not dies out so quickly since if the cluster head fails the entire cluster becomes useless and the member nodes are of no use .Thus a secondary cluster head is helpful in this case. Initially sink has all information about the number of nodes, energy levels of all nodes and location information of these nodes. Sink could get the whole location and energy information about sensor nodes by two or other manners. One is recorded in the sink at the initial state as nodes were deployed. The other is that sink broadcast whole network, and then received the back message form sensor nodes. The number of clusters in a network is predefined. So the number of cluster heads is also fixed.

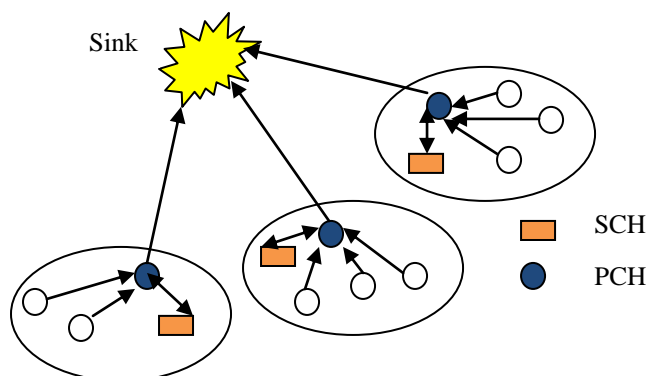


Figure 4. Proposed model

#### A. CLUSTER HEAD SELECTION

Based on the location information of the sensor nodes the cluster head is chosen .Initially the energy levels of every node is the same so for first round the cluster head is randomly selected .

#### B. PRIMARY CLUSTER HEAD SELECTION (PCH)

The primary cluster head is the main data aggregator in a cluster. For the first round the PCH is selected randomly based on the location of the nodes. The criteria for the selection of the PCH is that it should lie in the center of a cluster so that other member nodes can communicate with this PCH with minimum distance. From second round onwards the selection of PCH is based on the residual energy of nodes. The node with maximum residual energy in a cluster is nominated as the PCH by the base station.

#### C. SECONDARY CLUSTER HEAD SELECTION (SCH)

A SCH acts as a backup cluster head in a cluster. The criteria for the selection of the SCH are based on two factors:

1. Second maximum residual energy in a cluster.
2. Nearest node to the PCH.

Based on these factors, SCH is selected for each cluster. For the first round the second criteria is applicable only. Since, in the first round the energy of all nodes are the same.

#### D. SELECT CH ALGORITHM

1. For round = 1 :
  - a. PCH is selected randomly based on center of a cluster by sink
  - b. SCH is selected based on least distance from the PCH.
2. For round > 1 :
  - a. PCH is selected based on residual energy(max)
  - b. SCH is selected based on nearest to PCH and the second max residual energy

#### E. CLUSTER FORMATION

Once the Primary cluster head is selected then the cluster formation process starts. The PCH sends the broadcasts the join cluster msg in the network .the nodes that receive high signal strength from the PCH reply with an ACK msg to their respective PCH and thus join that cluster. When the cluster head receives the ACK message from its neighbor nodes , it assigns the node a time slot to transmit data based on TDMA schedule.

#### F. BUILD CLUSTER ALGORITHM

1. A selected PCH broadcasts a join cluster msg over neighbor nodes.
2. The neighboring nodes send a ACK msg to their neighboring PCH.
3. On receiving the ACK msg the PCH builds a cluster member list and a TDMA schedule and broadcasts them in the cluster.

#### G. ENTIRE OPERATION

After the cluster selection process completes, the sink node determines the member node of individual clusters with maximum residual energy other than the PCH node. That node is selected as the SCH node. Then each member node of a cluster sends data and its residual energy information to the cluster head by the given TDMA schedule. The cluster head maintains residual energy information of member nodes in a list and before the completion of a round, the cluster head sends base station this residual energy list. The sink node then determines the node with maximum residual energy and elects it as the PCH for the next round. The sink node also selects the



SCH based on the least distance from the PCH and the second maximum residual energy of nodes in a particular cluster. Now if during data transmission the PCH fails due to some reasons or it falls below the threshold energy then the entire responsibility of the data aggregation shifts onto the SCH. The SCH then informs about its new role to other member nodes and the data aggregation operation continues.

## V. CONCLUSION

In this work, we proposed a cluster based routing protocol that considers the residual energy of sensor nodes to extend the lifetime of sensor networks. Apart from this it includes an additional feature of a back up cluster head so that if the main cluster head fails due to energy depletion or any other reason then the responsibility of data aggregation in a particular cluster shifts to the back up cluster head. So this proposed new scheme is fault tolerant such that a cluster lifetime also is enhanced and is not dependent on a single cluster head.

## REFERENCES

- [1] Ian F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," IEEE Communications Magazine, Vol. 40, Issue 8, pp. 102-114, Aug. 2002.
- [2] K. Akkaya and M. Younis, "A Survey of Routing Protocols in Wireless Sensor Networks," in the Elsevier Ad Hoc Network Journal, Vol 3/3, pp. 325-349, 2005.
- [3] Q. Jiang and D. Manivannan, "Routing protocols for sensor networks," Proceedings of CCNC 2004, pp.93-98, Jan. 2004.
- [4] J. Ibric and I. mahgoub, "Cluster-Based Routing in Wireless Sensor Networks: Issues and Challenges", Proceedings of the 2004 Symposium on Performance Evaluation of Computer Telecommunication Systems, pp.759-766, Jul. 2004.
- [5] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," Proc. Hawaii Conf. System Sciences, Jan. 2000.
- [6] S. Lindsey, C. Raghavendra and K. M. Sivalingam, "Data Gathering Algorithms in Sensor Networks Using Energy Metrics," IEEE Transactions On Parallel and Distributed Systems, Vol . 13, No. 9, September 2002. PEGASIS
- [7] S. Lindsey, C. Raghavendra "Pegasis: Power-Efficient gathering in sensor information systems," In: Proc. of the IEEE Aerospace Conference, 2002. Pp.1-6.
- [8] V. Loscri, G. Morabito and S. Marano. "A Two-Levels Hierarchy for Low-Energy Adaptive Clustering Hierarchy".
- [9] Hang Zhou, Zhe Jiang and Mo Xiaoyan, "Study and Design on Cluster Routing Protocols of Wireless Sensor Networks", Dissertation, 2006.
- [10] Fan Xiangning, Song Yulin. "Improvement on LEACH Protocol of Wireless Sensor Network", 2007.